

SEQUENCE LISTING



RECEIVED

SEP 27 2000

TECH CENTER 1600/2900

<110> MENOZZI, Franco
LOCHT, Camille<120> IDENTIFICATION AND CLONING OF A MYCOBACTERIAL ANTIGEN
CORRESPONDING TO A HEPARIN-BINDING HAEMAGGLUTININ

<130> 960-34

<140> 09/192,579

<141> 1998-11-17

<150> FR 96 06168

<151> 1996-05-17

<160> 20

<170> PatentIn Ver. 2.1

RECEIVED

SEP 28 2000

TECH CENTER 1600/2900

<210> 1

<211> 39

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: peptide
sequence comprising a region involved in
interactions with sulphated glycoconjugates and in
heparin binding

<400> 1

Lys Lys Ala Ala Pro Ala Lys Lys Ala Ala Pro Ala Lys Lys Ala Ala
1 5 10 15Pro Ala Lys Lys Ala Ala Ala Lys Lys Ala Pro Ala Lys Lys Ala Ala
20 25 30Ala Lys Lys Val Thr Gln Lys
35

<210> 2

<211> 10

<212> PRT

<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: peptide S1441

<400> 2

Lys Ala Glu Gly Tyr Leu Glu Ala Ala Thr
1 5 10

<210> 3
<211> 18
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: peptide S1443

<400> 3
Xaa Glu Gly Tyr Val Asp Gln Ala Val Glu Leu Thr Gln Glu Ala Leu
1 5 10 15

Gly Lys

E
C
<210> 4
<211> 9
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: peptide S1446

<400> 4
Xaa Gln Glu Xaa Leu Pro Glu Xaa Leu
1 5

C
<210> 5
<211> 7
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: Peptide S1447

<400> 5
Phe Thr Ala Glu Glu Leu Arg
1 5

<210> 6
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
Oligonucleotide originated from the S1441 peptide
(oligo S1441)

<400> 6
aaggcsgagg gstacct

<210> 7
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
Oligonucleotide originated from the S1441 peptide
(reverse oligo S1441)

<400> 7
aggta~~sccctt~~ csgc~~ttt~~

17

<210> 8
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
Oligonucleotide originated from the S1443 peptide
(oligo S1443)

<400> 8
gaccaggc~~sg~~ tsgag~~c~~

17

E1
CY
<210> 9
<211> 17
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
Oligonucleotide originated from the S1443 peptide
(reverse oligo S1443)

<400> 9
agctcsac~~sg~~ cctgg~~t~~

17

<210> 10
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
Oligonucleotide named HBHASEq1 and used for
sequencing the gene coding for HBHA

<400> 10
agccggtaca accagctgg~~t~~ c

21

<210> 11
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
Oligonucleotide named HBHA Seq1inv and used for sequencing the gene coding for HBHA

<400> 11
gaccagctcg ttgtaccggc t

21

<210> 12
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
Oligonucleotide named HBHASEq2 and used for sequencing the gene coding for HBHA

<400> 12
catccaaacac gtcgactcc

19

E 1
<210> 13
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
Oligonucleotide named HBHA Seq3 and used for sequencing the gene coding for HBHA

<400> 13
ttgatgtcat caatgttcg

19

Q 2
<210> 14
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
Oligonucleotide named HBHA Seq4 and used for sequencing the gene coding for HBHA

<400> 14
cgtggaccag gcgggtggag

19

<210> 15
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
Oligonucleotide named HBHA Seq 5 and used for sequencing the gene coding for HBHA

<400> 15
gacgatcagg aggttccccc g

21

<210> 16
<211> 24
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence:
Oligonucleotide named reverse primer and used for sequencing the gene coding for HBHA

<400> 16
agcggtataac aatttcacac agga

24

E1
~~Q1~~
<210> 17
<211> 149
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: nucleotide sequence and amino sequence of a fragment of HBHA deduced from a PCR fragment of chromosomal BCG DNA

<220>
<221> CDS
<222> (1)...(147)

<400> 17
aag gcc gag ggc tac ctc gag gcc gcg act agc cg^g tac aac gag ctg
Lys Ala Glu Gly Tyr Leu Glu Ala Ala Thr Ser Arg Tyr Asn Glu Leu
1 5 10 15

48

g^tc gag cgc ggt gag gcc gct cta gag cg^g ctg cgc agc cag cag agc
Val Glu Arg Gly Glu Ala Ala Leu Glu Arg Leu Arg Ser Gln Gln Ser
20 25 30

96

t^tc gag gaa gt^t tcg gc^g ccc gcc gaa gg^c tac gt^t gac cag gc^g gtc
Phe Glu Glu Val Ser Ala Pro Ala Glu Gly Tyr Val Asp Gln Ala Val
35 40 45

144

gag ct
Glu

149

<210> 18
<211> 49
<212> PRT
<213> Artificial Sequence
<223> Description of Artificial Sequence: amino sequence of a fragment of HBHA deduced from a PCR fragment of chromosomal BCG DNA

<400> 18
Lys Ala Glu Gly Tyr Leu Glu Ala Ala Thr Ser Arg Tyr Asn Glu Leu
1 5 10 15

Val Glu Arg Gly Glu Ala Ala Leu Glu Arg Leu Arg Ser Gln Gln Ser
20 25 30

Phe Glu Glu Val Ser Ala Pro Ala Glu Gly Tyr Val Asp Gln Ala Val
35 40 45

Glu

E1
<210> 19
<211> 1097
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: DNA sequence of the BCG gene coding for HBHA

<220>
<221> CDS
<222> (331)..(924)
<223> CDS from 811 to 828, from 829 to 846, from 847 to 864, from 865 to 885 and from 895 to 915 : peptide which may be particularly involved in interactions with sulphated glycoconjugates

<400> 19
cggtggcg gtaatcaaac ctgaaggaca gtcatctggg tgaggtcgac cgcaggctga 60
tccagccgat cggccggcgc tggccaacag cgactccgtc gatgacgtgc agcaaaggag 120
acatgtatgt accggatcag ctggccctga catctacgaa ctcgaccgac aaccgaccgg 180
acgatcagga gttttccccg gcaagtgcgc tgccatgtca atccgcgggt ctgactatgt 240
cctccctgga ggagccgacg cttgccccaa cgtccagacc aaagatgtaa gaacgcccgt 300
atcagaaaaat agttaatgaa aggaataccc atg gct gaa aac tcg aac att gat 354
Met Ala Glu Asn Ser Asn Ile Asp
1 5

gac atc aag gct ccg ttg ctt gcc gcg ctt gga gcg gcc gac ctg gcc 402
Asp Ile Lys Ala Pro Leu Leu Ala Ala Leu Gly Ala Ala Asp Leu Ala
10 15 20

ttg gcc act gtc aac gag ttg atc acg aac ctg cgt gag cgt gcg gag 450
 Leu Ala Thr Val Asn Glu Leu Ile Thr Asn Leu Arg Glu Arg Ala Glu
 25 30 35 40

gag act cgt acg gac acc cgc agc cgg gtc gag gag agc cgt gct cgc 498
 Glu Thr Arg Thr Asp Thr Arg Ser Arg Val Glu Glu Ser Arg Ala Arg
 45 50 55

ctg acc aag ctg cag gaa gat ctg ccc gag cag ctc acc gag ctg cgt 546
 Leu Thr Lys Leu Gln Glu Asp Leu Pro Glu Gln Leu Thr Glu Leu Arg
 60 65 70

gag aag ttc acc gcc gag gag ctg cgt aag gcc gac ggc tac ctc 594
 Glu Lys Phe Thr Ala Glu Leu Arg Lys Ala Ala Glu Gly Tyr Leu
 75 80 85

gag gcc gcg act agc cgg tac aac gag ctg gtc gag cgc ggt gag gcc 642
 Glu Ala Ala Thr Ser Arg Tyr Asn Glu Leu Val Glu Arg Gly Glu Ala
 90 95 100

gct cta gag cgg ctg cgc agc cag cag agc ttc gag gaa gtg tcg cgc 690
 Ala Leu Glu Arg Leu Arg Ser Gln Gln Ser Phe Glu Glu Val Ser Ala
 105 110 115 120

ccc gcc gaa ggc tac gtg gac cag cgc gtg gag ttg acc cag gag gcg 738
 Pro Ala Glu Gly Tyr Val Asp Gln Ala Val Glu Leu Thr Gln Glu Ala
 125 130 135

ttg ggt acg gtc gca tcg cag acc cgc gcg gtc ggt gag cgt gcc gcc 786
 Leu Gly Thr Val Ala Ser Gln Thr Arg Ala Val Gly Glu Arg Ala Ala
 140 145 150

aag ctg gtc ggc atc gag ctg cct aag aag gct gct ccg gcc aag aag 834
 Lys Leu Val Gly Ile Glu Leu Pro Lys Lys Ala Ala Pro Ala Lys Lys
 155 160 165

gcc gct ccg gcc aag aag gcc gct ccg gcc aag aag gcg gcg gcc aag 882
 Ala Ala Pro Ala Lys Lys Ala Ala Pro Ala Lys Lys Ala Ala Ala Lys
 170 175 180

aag gcg ccc gcg aag aag gcg gcg gcc aag aag gtc acc cag 924
 Lys Ala Pro Ala Lys Lys Ala Ala Lys Lys Val Thr Gln
 185 190 195

aagttagtcgg gctccgaatc accatcgact ccgagtcgcc cacggggcga ctcggagtgc 984
 acgtgttgg a tgcaaaccgc atagtcgaa tgcgtgagcc acctcgtggg taccgtcatg 1044
 ctggattgc tggtcgcccgt ctgggtgaca gcgggttacg cgtttgca tgc 1097

<210> 20
 <211> 198
 <212> PRT
 <213> Artificial Sequence
 <223> Description of Artificial Sequence: DNA sequence
 of the BCG gene coding for HBHA

<400> 20
 Met Ala Glu Asn Ser Asn Ile Asp Asp Ile Lys Ala Pro Leu Leu Ala
 1 5 10 15

Ala Leu Gly Ala Ala Asp Leu Ala Leu Ala Thr Val Asn Glu Leu Ile
20 25 30

Thr Asn Leu Arg Glu Arg Ala Glu Glu Thr Arg Thr Asp Thr Arg Ser
35 40 45

Arg Val Glu Glu Ser Arg Ala Arg Leu Thr Lys Leu Gln Glu Asp Leu
50 55 60

Pro Glu Gln Leu Thr Glu Leu Arg Glu Lys Phe Thr Ala Glu Glu Leu
65 70 75 80

Arg Lys Ala Ala Glu Gly Tyr Leu Glu Ala Ala Thr Ser Arg Tyr Asn
85 90 95

Glu Leu Val Glu Arg Gly Glu Ala Ala Leu Glu Arg Leu Arg Ser Gln
100 105 110

Gln Ser Phe Glu Glu Val Ser Ala Pro Ala Glu Gly Tyr Val Asp Gln
115 120 125

Ala Val Glu Leu Thr Gln Glu Ala Leu Gly Thr Val Ala Ser Gln Thr
130 135 140

Arg Ala Val Gly Glu Arg Ala Ala Lys Leu Val Gly Ile Glu Leu Pro
145 150 155 160

Lys Lys Ala Ala Pro Ala Lys Lys Ala Ala Pro Ala Lys Lys Ala Ala
165 170 175

Pro Ala Lys Lys Ala Ala Ala Lys Lys Ala Pro Ala Lys Lys Ala Ala
180 185 190

Ala Lys Lys Val Thr Gln
195